



HAL
open science

Évaluations ergonomiques d'interfaces humain-machines pour la défense : un exemple de système de surveillance maritime

Pierre Laborde, Éric Le Pors, Yann Le Goff, Alain Plantec, Steven Costiou

► To cite this version:

Pierre Laborde, Éric Le Pors, Yann Le Goff, Alain Plantec, Steven Costiou. Évaluations ergonomiques d'interfaces humain-machines pour la défense : un exemple de système de surveillance maritime. IHM'24 - 35e Conférence Internationale Francophone sur l'Interaction Humain-Machine, AFIHM; Sorbonne Université, Mar 2024, Paris, France. hal-04498410

HAL Id: hal-04498410

<https://hal.science/hal-04498410>

Submitted on 11 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Évaluations ergonomiques d'interfaces humain-machines pour la défense : un exemple de système de surveillance maritime

Ergonomic evaluations of Human-Machine Interfaces in the defense business: an example of a collaborative Maritime Surveillance system

PIERRE LABORDE^{**}, THALES Defense Mission Systems France, France

ÉRIC LE PORS^{*}, THALES Defense Mission Systems France, France

YANN LE GOFF[†], Univ Brest, CNRS, Lab-STICC, CS 93837, France

ALAIN PLANTEC[†], Univ Brest, CNRS, Lab-STICC, CS 93837, France

STEVEN COSTIOU[‡], Inria, Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISTAL, France

Chez *Thales Defense Mission System*, nous impliquons les utilisateurs finaux dans la conception des systèmes logiciels, en particulier lors d'évaluations de prototypes d'Interfaces Humain-Machine. Dans les métiers de la défense, les utilisateurs finaux disposent de peu de disponibilités pour participer à ces évaluations. Leurs visites sont concentrées sur quelques jours et espacées de plusieurs mois ou années. Pour maximiser les retours, nous organisons des sessions d'évaluations ergonomiques sur des scénarios longs et complexes pour confronter nos IHM à un panorama exhaustif des situations que les utilisateurs peuvent rencontrer lors de leurs missions. Ces évaluations se déroulant sur des prototypes, il arrive que des événements imprévus perturbent le déroulement des scénarios exécutés, par exemple des erreurs logicielles.

Dans cet article, nous décrivons notre approche pour faire évoluer dynamiquement nos prototypes en cours d'évaluation et bénéficier au maximum de la présence de nos utilisateurs finaux sans interrompre les scénarios joués.

CCS Concepts: • **Human-centered computing** → **HCI design and evaluation methods**; • **Software and its engineering** → **Software prototyping**.

Mots Clés et Phrases Supplémentaires: Évaluations ergonomiques, IHM pour les systèmes de la défense, patrouille maritime, live programming

Additional Key Words and Phrases: Ergonomic evaluations, HCI for defense systems, maritime patrol, live programming

Reference:

Pierre Laborde, Éric le Pors, Yann Le Goff, Alain Plantec, and Steven Costiou. 2024. Évaluations ergonomiques d'interfaces humain-machines pour la défense : un exemple de système de surveillance maritime. *IHM '24 : Articles Industriels de la 35^e conférence Francophone sur l'Interaction Humain-Machine, March 25–29, 2024, Paris, France*

1 INTRODUCTION

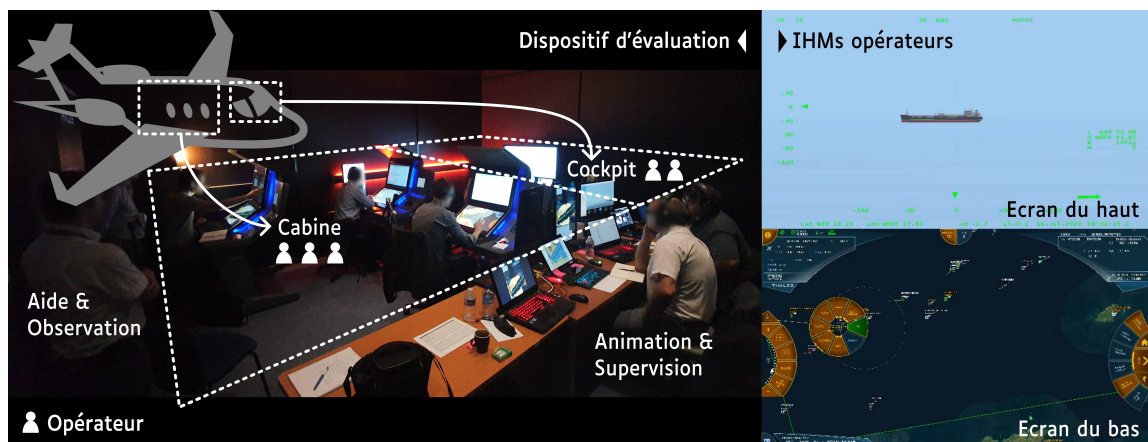
Thales Defense Mission Systems (DMS) développe des logiciels pour le milieu de la défense. Ces logiciels subissent de longs et coûteux procédés de conception et de validation, notamment de leurs Interfaces Humain-Machine (IHM). Une fois réalisées, il est extrêmement coûteux de revenir sur la conception des IHM. Nos décisions de conception IHM sont donc guidées par des évaluations ergonomiques effectuées en amont du développement du logiciel final. Ces évaluations nécessitent de mettre des opérateurs (nos utilisateurs finaux), tels que ceux de la surveillance maritime, dans des contextes et des situations les plus représentatives de la réalité possible. Ces opérateurs sont très peu disponibles et ne peuvent participer à des évaluations que quelques jours par an. En dehors de la présence des opérateurs en évaluation,

*. Both authors contributed equally to this research.

il est difficile d'avoir un retour rigoureux sur la conception et ses évolutions. Pendant ces évaluations, nous sommes contraints à capturer un maximum de retours et à pouvoir faire évoluer la conception en cours d'évaluation.

Pour pallier ces problèmes, nous développons et maintenons des prototypes afin d'évaluer nos IHM au travers de scénarios d'usage réalistes. Nous matérialisons toutes les contraintes et situations dans lesquelles les opérateurs peuvent physiquement se retrouver au cours d'une mission réelle (par exemple, l'intérieur d'un avion de surveillance maritime). Lorsqu'un problème est détecté (bug, problème ergonomique, fonctionnalité manquante, etc..) nous intervenons en temps réel, sans interrompre l'évaluation, afin de résoudre le problème et permettre une correction immédiate. Nous constatons que nos opérateurs sont généralement satisfaits, acceptent et adoptent les IHM modifiées suite à leurs remarques et font moins de demandes de correction a posteriori. Dans cet article, nous prenons pour exemple l'évaluation d'un système de surveillance maritime. Nous présentons notre méthode pour dynamiquement appliquer des évolutions dans le système en cours d'exécution et les évaluer immédiatement avec les opérateurs.

2 EXEMPLE : DISPOSITIF D'ÉVALUATION ERGONOMIQUE EN SURVEILLANCE MARITIME



Équipage de surveillance maritime de 5 opérateurs installés à bord d'un avion. Le pilote et le chef de mission sont dans le cockpit. Le pilote manipule un simulateur de pilotage. Le chef de mission dispose d'une tablette avec la situation tactique. Les 3 opérateurs de la cabine ont chacun deux écrans tactiles, une boule roulante et un clavier. Deux écrans sur les côtés (hublots) et un grand écran à l'avant affichent des représentations 3D réalistes de l'environnement extérieur de l'avion.

Fig. 1. Un dispositif d'évaluation ergonomique en surveillance maritime, ainsi que l'une des IHM évaluée.

La Figure 1 présente un exemple type d'évaluation ergonomique [6]. Dans cet exemple, nous voyons l'installation d'un système de surveillance maritime lors d'une séance d'évaluation ergonomique avec un équipage de 5 opérateurs. Une cellule d'aide et d'observation complète le dispositif avec des ergonomes et des ingénieurs spécialisés en prototypage pour repérer et corriger en direct les bugs qui peuvent apparaître durant l'évaluation ergonomique. Les ergonomes prennent des notes et n'interviennent que sur demande d'aide explicite des opérateurs.

Une cellule de supervision et d'animation contrôle le bon déroulement des évaluations. Elle peut techniquement activer des parties spécifiques du scénario et modifier le comportement des acteurs, par exemple ajouter un navire inconnu dans la zone de mission. Dans cette cellule, des ingénieurs et des anciens opérateurs sont présents pour rendre le scénario vivant, pour l'animer : ils jouent un jeu de rôle avec l'équipage. Un des superviseurs dispose d'un poste

avec une duplication des IHMs des opérateurs. Il est chargé des modifications logicielles qui peuvent être déployées dynamiquement, sans redémarrer, sur l'ensemble des postes opérateurs.

Nous réalisons ces évaluations ergonomiques pour mesurer l'efficacité et l'efficience des IHMs d'un système [2], ainsi que le ressenti des utilisateurs. Ces mesures nous permettent aussi d'avoir une idée de l'intuitivité des IHMs. Ces évaluations ont toujours lieu en présence des opérateurs. Les opérateurs utilisent les prototypes IHM et interagissent avec le système comme lors d'une vraie mission. Nous devons rendre l'évaluation la plus proche possible de la réalité pour améliorer la qualité de l'évaluation, la pertinence des remarques et être sûr d'adresser les bons problèmes. Les scénarios sont basés sur des données AIS (*Automatic Identification System*) réelles du trafic maritime avec plusieurs centaines de bateaux et des éléments aléatoires superposés aux données réelles pour simuler les imprévus.

Ainsi, nous reproduisons les facteurs que nous estimons les plus impactants pour les opérateurs tels que l'ergonomie du poste opérateur mais aussi les contraintes physiques et psychologiques qui contribuent à se rapprocher des conditions réelles de mission (ambiance sonore, charge de travail, charge cognitive, situation inconnue ou imprévue, etc.).

3 PROBLÈME : LA DISPONIBILITÉ DES OPÉRATEURS POUR DES ÉVALUATIONS DE LONGUE DURÉE

Pour évaluer nos IHM, nous procédons à des exécutions de scénarios inspirés de missions du domaine cible réellement menées. Les actuels opérateurs de ces missions sont les futurs utilisateurs des logiciels dont les prototypes IHM sont évalués. L'activité des opérateurs en cours de mission peut être très complexe. Les missions sont de longue durée (plusieurs heures) et sont effectuées sans interruption. Certaines périodes plus critiques de la mission nécessitent beaucoup de concentration. Les opérateurs peuvent cumuler fatigue et voir leurs capacités cognitives réduites.

Les scénarios des évaluations ergonomiques et leur exécution doivent être réalistes et reproduire le plus fidèlement possible les conditions réelles des missions. Les séances d'évaluation ergonomiques doivent donc pouvoir durer plusieurs heures sans interruption. Dans notre exemple, les opérateurs exécutent des scénarios de 3 heures sur une période de 3 jours (soit 9h en tout), à cela s'ajoute les temps de préparation (briefing) et de restitution.

Cependant, les opérateurs sont peu disponibles. Dans le contexte du projet industriel de l'exemple de la Figure 1, les opérateurs n'ont été présents que sur deux périodes de deux à trois jours sur une durée totale de 8 années de projet. Il n'est pas rare que la présence des opérateurs ne se limite qu'à une ou deux périodes de quelques jours sur toute la phase de conception. Sur d'autres projets, les opérateurs n'ont l'opportunité de ne participer aux évaluations qu'une seule fois sur des périodes de plusieurs mois. Il nous est alors indispensable de maximiser les retours d'expérience pendant la présence des opérateurs. Or, une évaluation ergonomique réussie consiste :

- en l'exécution d'un ou plusieurs scénarios fidèles à la réalité, en utilisant les prototypes sans interruption ;
- en un bilan ergonomique positif émanant du retour direct des opérateurs (futurs utilisateurs).

Dans ce contexte, étant donné la durée d'une évaluation et la durée de présence limitée des opérateurs, les problèmes majeurs sont le risque d'interruption d'une exécution dû à une modification devant être apporté au prototype lui même et la nécessité d'adapter une partie d'un scénario pour qu'il demeure fidèle à la réalité. Une interruption de l'exécution d'un scénario peut invalider l'évaluation en cours et implique le redémarrage du scénario depuis son début. Il nous est actuellement impossible de reproduire les heures d'interactions des opérateurs avec leur poste et les interactions collectives entre tout ou partie des opérateurs. Par exemple, les séries d'évènements issues des manipulations du matériel sont impossibles à reproduire. Chaque exécution de scénario consiste en un contexte unique et les mesures ergonomiques effectuées avant un redémarrage imprévu deviennent inexploitable.

Typiquement, deux raisons peuvent conduire à l'interruption d'une exécution : un dysfonctionnement ou une demande de modification d'une IHM. Les opérateurs peuvent en effet demander des corrections ou des adaptations des IHMs pendant l'exécution d'un scénario, si ces dernières présentent des défauts ergonomiques.

4 NOTRE APPROCHE : ADAPTATION DIRECTE DU SYSTÈME EN COURS D'ÉVALUATION

Pour maintenir la fluidité d'exécution de nos scénarios, nous explorons et corrigeons les problèmes imprévus directement pendant l'exécution d'un scénario, sans interrompre l'évaluation en cours. Pour cela, nous intervenons dynamiquement sur le prototype afin de l'adapter.

4.1 Intervention à chaud sur le système en cours d'évaluation

Nous modifions le système en cours d'évaluation de deux manières : en modifiant dynamiquement le code des prototypes et en intervenant directement dans les scénarios joués.

Modification dynamique de code. Nous réalisons nos prototypes avec notre atelier logiciel de prototypage *Smalltalk Mockup (Smock)*. *Smock* est écrit dans différents dialectes de *Smalltalk (VisualWorks [1] et Pharo [5])*. Nous avons choisi *Smalltalk* pour ses propriétés de *Live Programming [3, 7]*. Ces propriétés nous permettent, sans interrompre l'exécution d'un prototype, de surveiller dynamiquement son état, d'éditer et de recompiler son code ainsi qu'à appliquer et observer instantanément l'impact de nos modifications.

Lorsqu'un opérateur rencontre un problème, nous modifions le code à l'aide du debugger sans que les autres opérateurs n'en soient conscients et sans que le reste de l'évaluation ne soit affecté, c'est-à-dire sans arrêter ou redémarrer les prototypes en cours d'évaluation. La modification dynamique du code nous donne la possibilité de comprendre les contraintes dans un contexte donné et d'expérimenter des alternatives de code sans perdre ou invalider le contexte, les données et les mesures d'une évaluation en cours.

Avant de choisir *Smalltalk*, nous avons évalué d'autres langages et technologies de mise en oeuvre comme *Java*, *C#* ou *Objective-C*. Parmi ces technologies, seul *Smalltalk* expose les propriétés que nous recherchions pour la modification dynamique de nos prototypes, avec un environnement de développement construit et adapté pour le live programming.

Intervention d'acteurs humains. Certains problèmes ne peuvent pas être immédiatement résolus de manière technique. Pour maintenir la fluidité de l'exécution d'un scénario, nous utilisons la technique dite du *Magicien d'Oz [4]*. Nous intervenons pour simuler des événements réels avec des façades qui donnent l'impression aux opérateurs d'interagir avec un vrai logiciel ou une vraie situation. Par exemple dans l'évaluation de la Figure 1, nous utilisons au cours d'un scénario une illustration dessinée à la main que nous présentons par le hublot pour simuler un événement réel. S'agissant d'évaluer le comportement des opérateurs qui exécutent un scénario, cette technique est suffisante dans le cadre de nos évaluations.

4.2 Exemples d'intervention sur un système en cours d'évaluation

Nous présentons deux exemples réels, dans le cadre de l'évaluation illustrée par la Figure 1 : une demande de modification d'une vue par un opérateur et la correction à chaud de bugs.

Un opérateur veut modifier une vue. Le problème rencontré durant une évaluation est le suivant : la tranche avant (pilote et chef de bord) n'avait pas certaines informations sur les bateaux détectés par le système de mission. La tranche arrière (cabine) avait ces informations. La tranche avant passait donc beaucoup de temps à demander ces informations

à tout l'équipage sur le canal vocal. En découlait un délai de réponse, une attention supplémentaire de la tranche arrière pour trouver les informations, un effort de coordination de la tranche arrière pour savoir qui donnerait l'information, et une forte occupation du canal vocal.

Le chef de bord a donc reporté à un superviseur que ces informations pourraient être affichées dans une des IHM, ce qui lui éviterait de les demander à la tranche arrière. Le superviseur réalise la correction à chaud, sans interrompre le scénario, puis informe le chef de bord qu'il peut de nouveau ouvrir la vue. Le chef de bord ouvre la vue, constate la modification et est satisfait du changement. L'équipage a alors arrêté d'utiliser le canal vocal pour communiquer cette information puisque qu'elle était désormais disponible sur l'IHM du chef de bord. L'ajout dynamique d'une information supplémentaire sur l'écran d'un seul opérateur a libéré le canal audio de tout l'équipage.

Nous avons pu noter de manière générale que la modification des IHM pour présenter des informations demandées par les opérateurs entraîne une baisse de l'utilisation des communications vocales, ce qui modifie l'intensité de l'utilisation du canal audio et la nature des informations qui y sont échangées. Le canal audio devient beaucoup moins saturé et les informations qui y circulent de plus forte importance que précédemment.

Modification du code à la volée pour corriger des bugs. Lorsqu'un opérateur rencontre un bug, nous devons intervenir sans interrompre et sans redémarrer le scénario. Voici plusieurs exemples de l'évaluation illustrée par la Figure 1 :

- Un potentiel bug est remonté sur une IHM. Nous patchons à la volée cette IHM sans que les opérateurs ne le sachent, et nous déployons la modification sur tous les postes.
- Un opérateur rencontre un bug sur une IHM, et un debugger s'ouvre sur son poste. Nous intervenons directement sur le poste opérateur problématique. Nous utilisons le debugger pour comprendre et corriger le problème directement à partir du debugger. Seul le poste opérateur concerné est affecté, et le debugging est transparent pour les autres postes qui continuent le scénario. Par la suite la correction est déployée sur les autres postes.
- Dans certains cas, un bug peut ne pas interrompre l'usage d'une IHM, le scénario continue pour le poste opérateur en question. Le superviseur utilise sa propre instance indépendante pour comprendre le bug, le corriger et déployer sa correction sur tous les postes opérateurs.

Le live programming permet d'obtenir des informations contextuelles n'apparaissant qu'après un certain temps de scénario, ou qu'après une séquence précise et parfois imprévue d'actions des opérateurs. Ces bugs ne pourraient pas facilement être reproduits dans un environnement de test sans opérateurs.

Pour conserver la fluidité des scénarios joués, nous devons minimiser le temps passé à debugger. Nous nous fixons pour cela une limite empirique à 3 minutes. Par exemple, sur une évaluation ergonomique de 3 jours au cours de laquelle nous avons joué 4 scénarios, nous avons rencontré 5 bugs. Tous les bugs se sont manifestés tard dans les scénarios : à 54% de progression pour le bug apparu le plus tôt et à 97% de progression pour le plus tardif. Dans cette situation, il est impensable de redémarrer l'évaluation entièrement. Les observateurs sont parvenus à corriger tous les bugs autour de la limite de 3 minutes, en 1 minute 25 secondes pour le plus court et en 4 minutes 27 secondes pour le plus long. Ces délais d'interruptions sont suffisamment courts pour conserver la fluidité des scénarios.

5 ÉVALUATION DE L'APPROCHE

Le but d'une évaluation ergonomique est de valider avec un client (le demandeur) et les opérateurs (les utilisateurs) un ensemble de points critiques afin de stabiliser le périmètre fonctionnel ainsi que les coûts de développement logiciel afférents.

Avantages de nos évaluations ergonomiques. Une phase d'évaluation ergonomique utilisant notre méthode de prototype a de nombreux avantages.

Les retours des utilisateurs prennent habituellement des semaines avant d'être pris en compte. Il faut livrer une version et attendre les retours. Pendant ce temps le client utilise un produit qui génère des irritants voire un doute dans la capacité de son système. Utiliser un prototype dynamiquement modifié pour satisfaire les demandes du client (via les opérateurs lors des évaluations) permet de limiter ces phases de livraison-retour et de frustration lors de l'utilisation d'un produit inadéquat. La démarche réduit notamment le coût de retravail des IHM habituellement effectué au fil de l'eau. Les remarques des opérateurs lors des évaluations et les évolutions qui en découlent démontrent un fort retour sur investissement.

Les évaluations nous permettent également de vérifier que le prototype remplit le contour fonctionnel attendu en couvrant le déroulement de différents scénarios d'emploi. Les évaluations sur prototype permettent également de s'assurer de la cohérence de l'ergonomie et du design graphique. La navigation libre des opérateurs d'un bout à l'autre de l'IHM permet de valider empiriquement son utilisabilité. De plus, Le prototype permet de valider la dynamique des données et comment l'opérateur les voit, les utilise ou les enrichit. Cela nous permet de comprendre comment l'opérateur navigue en utilisant l'IHM pour réaliser ces actions (en matérialisation les chemins utilisateur).

Limites. Néanmoins, ce processus reste limité. Afin de bénéficier d'un maximum de retours et de validation des opérationnels, il est nécessaire de commencer la phase de prototypage au moment où les différentes ingénieries sont en accord sur les modèles de données et les scénarios d'emploi du système à évaluer. Aujourd'hui, il reste difficile de convaincre d'un retour sur investissement en amont des domaines et métiers de notre entreprise même si des retours existent dans d'autres domaines.

Les évaluations se limitent souvent aux cas nominaux identifiés avec le client en phase de travail amont. Un prototype, même complexe, ne permet pas de tester le système dans ses limites de performances maximale, même si nous testons 80% du cas d'usage nominal. Cela nous permet de tester une grande partie de la charge de travail des opérateurs.

Impact des interventions directes sur le système en cours d'évaluation. Des évolutions IHM sont demandées et directement appliquées en évaluations. L'impact sur l'efficacité et l'efficience est visible grâce à l'analyse des mesures et des notes. Nous avons déjà observé une baisse de la charge mentale et une augmentation globale de l'efficacité de l'équipage. Les équipes d'ingénierie système capturent les évolutions des concepts d'emploi et d'usages opérationnels suite à ces demandes et voient des retours concrets. Sans évaluation ergonomique avec un prototype, toutes ces évolutions ne seraient probablement pas détectés avant la mise en opération du système.

6 CONCLUSION

Nous avons présenté comment, chez *Thales Defense Mission System*, nous impliquons nos utilisateurs finaux (ou opérateurs) dans l'évaluation de nos prototypes d'interfaces humain-machine pour les systèmes de mission de défense. Cependant, nos évaluations sont longues et nos utilisateurs finaux peu disponibles. Nous utilisons une approche permettant d'ajuster dynamiquement nos prototypes pendant une évaluation afin de tirer pleinement parti de la présence des opérateurs sans interrompre les scénarios ou perturber la collecte de données.

RÉFÉRENCES

- [1] [n.d.]. CINCOM Visual Works. <https://www.cincomsmalltalk.com/main/products/visualworks/>. Accessed : 2024-01-18.
- [2] [n.d.]. ISO 9241-210 :2019 Ergonomie de l'interaction homme-système. Partie 210 : Conception centrée sur l'opérateur humain pour les systèmes interactifs. <https://www.iso.org/fr/standard/77520.html>. Accessed : 2024-01-19.
- [3] Ademar Aguiar, André Restivo, Filipe Figueiredo Correia, Hugo Sereno Ferreira, and João Pedro Dias. 2019. Live software development : Tightening the feedback loops. In *Companion Proceedings of the 3rd International Conference on the Art, Science, and Engineering of Programming*. 1–6.
- [4] René Amalberti and Claude Valot. 1992. Le Magicien d'Oz. *CERMA, journée du PRC Rhône-Alpes* (1992).

- [5] Andrew P. Black, Stéphane Ducasse, Oscar Nierstrasz, Damien Pollet, Damien Cassou, and Marcus Denker. 2009. *Pharo by Example*. Square Bracket Associates, Kehrsatz, Switzerland. 333 pages. <http://books.pharo.org>
- [6] Catherine Letondal, Pierre-Yves Pillain, Emile Verdurand, Daniel Prun, and Olivier Grisvard. 2014. Of Models, Rationales and Prototypes : Studying Designer Needs in an Airborne Maritime Surveillance Drawing Tool to Support Audio Communication. In *HCI 2014, 28th BCS International Conference on Human-Computer Interaction*. pp–8.
- [7] David Ungar, Henry Lieberman, and Christopher Fry. 1997. Debugging and the experience of immediacy. *Commun. ACM* 40, 4 (1997), 38–43.